

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-129572

(43)Date of publication of application : 19.05.1995

(51)Int.Cl.

G06F 17/22

(21)Application number : 05-274604

(71)Applicant : MATSUSHITA ELECTRIC IND CO LTD

(22)Date of filing : 02.11.1993

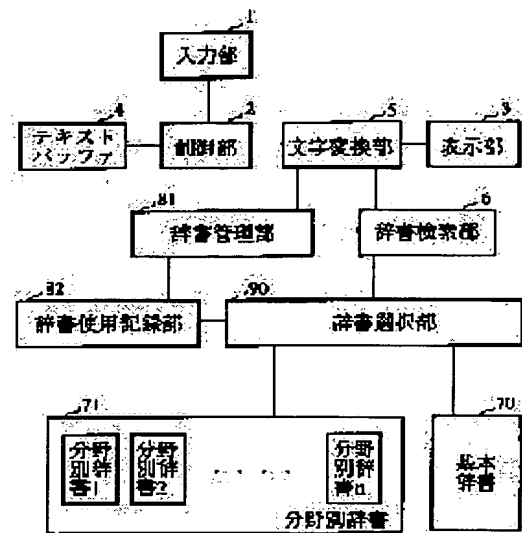
(72)Inventor : KAKU TOSHITAKE

(54) KANJI CONVERTER

(57)Abstract:

PURPOSE: To select a proper dictionary in a KANJI (Chinese character) converter provided with a basic dictionary and plural field dictionaries.

CONSTITUTION: When phonetic symbol string is inputted, an input part 1 segments the reading code of a morpheme corresponding to the phonetic symbol to sent to a dictionary retrieving part 6. A dictionary usage recording part 82 records the using time of each dictionary. A dictionary selection part 90 decides the retrieving order by giving priority to the dictionary of more using times in the respective dictionaries 70 and 71, which is recorded in the dictionary usage recording part 82. A dictionary retrieving part 6 retrieves the inside of each dictionary corresponding to the priority order by setting the reading code sent from the input part 1 as a retrieving key and detects all the corresponding homophones and word candidates. A character conversion part 5 makes a display part to display the detected homophones and word candidates, makes a user to confirm right conversion and sends the code of a dictionary registering the homophone and word which are judged to be right to a dictionary managing part 81. The dictionary managing part 81 updates the dictionary usage recording part 82.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision]

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平7-129572

(43)公開日 平成7年(1995)5月19日

(51)Int.Cl. ⁶	識別記号	序内整理番号	F I	技術表示箇所
G 0 6 F 17/22		7315-5L	G 0 6 F 15/ 20	5 2 4 A
		7315-5L		5 2 0 G

審査請求 未請求 請求項の数3 O L (全 9 頁)

(21)出願番号 特願平5-274604

(22)出願日 平成5年(1993)11月2日

(71)出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72)発明者 郭 俊桔

台湾タイ・ペイ・シ・ター・アン・チー・

10628・レン・アイ・ル・サン・トォア

ン・136・ハオ・ロウ スン・シャ・ティ

エン・チ・チ・シュー・カイ・ファー・ク

ウー・フェン・ユウ・シエン・コン・スー

内

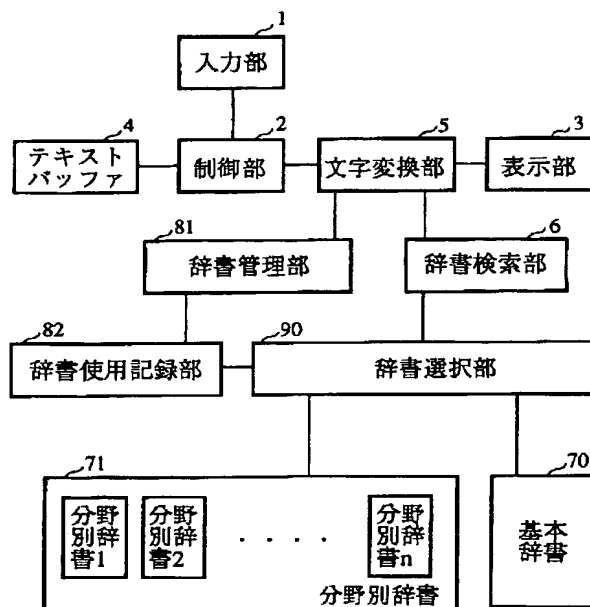
(74)代理人 弁理士 中島 司朗

(54)【発明の名称】 漢字変換装置

(57)【要約】

【目的】 基本辞書及び複数の分野別辞書を備えた漢字変換装置において、適切な辞書を選択する。

【構成】 表音文字列が入力されると、入力部はこれに相応する形態素の読みコードを切り出して辞書検索部へ送る。辞書使用記録部は、各辞書毎にその使用回数を記録している。辞書選択部は、辞書使用記録部に記録されている各辞書の使用回数の多いものを優先して、その検索順序を決める。辞書検索部は、入力部から送られてきた読み記号を検索キーとして優先順位に従って各辞書内の検索を行ない、対応する全ての同音字、語候補を検出する。文字変換部は、検出された同音字、語候補を表示部に表示させ、使用者に正しい変換を確認させ、併せて正しいとされた同音字、語の登録されている辞書のコードを辞書管理部へ送る。辞書管理部は、これにより、辞書使用記録部を更新する。



【特許請求の範囲】

【請求項 1】 入力された表音文字列に相応する形態素の読みコードを切り出して辞書検索部へ送る入力部と基本辞書と複数の分野別辞書とからなる辞書部と表示部とを有する漢字変換装置において、

各辞書についてその使用回数を記録する辞書使用記録部と、

上記辞書使用記録部に記録されている使用回数を多いものを優先して各辞書の検索順序を決める辞書選択部と、上記入力部により送られてきた読み記号を検索キーとして、上記辞書選択部が決める検索順序に従って各辞書の検索を行ない、読み記号に対応する全ての同音字、語候補及び当該同音字、語候補毎にそれらが登録されている分野別辞書のコードを取り出して文字変換部へ送る辞書検索部と、

上記辞書検索部により検索された同音字、語の候補を表示部に表示させ使用者に正しい変換を確認させると共に、正しいとされた同音字、語の登録されている辞書のコードを辞書管理部へ送る文字変換部と、

上記文字変換部から送られてきた辞書コードをもとに、上記辞書使用記録部の記録する各辞書の使用回数を所定の手順で更新する辞書管理部とを有していることを特徴とする漢字変換装置。

【請求項 2】 前記辞書管理部は、上記辞書使用記録部の記録を更新する際に、使用回数が記録されている複数の辞書のうち、所定数のものの使用回数があらかじめ定められたしきい値に達したときには、それまでの若しくは一まとまりの漢字変換の終了したときまでの使用回数がそのしきい値に達していない辞書に対してはその使用回数から一定値を引き去るか 0 にし、それまでの使用回数が 0 か、前記引き去りの結果負となった辞書に対しては不使用に設定することを特徴とする請求項 1 記載の漢字変換装置。

【請求項 3】 前記表示部は、前記辞書検索部が所定数の辞書で読み記号に対応する同音字、語候補の検索が成功した場合には、その段階でその検索に成功した同音字、語候補のみを使用者による正しい変換の確認のために表示部に表示することを特徴とする請求項 1 若しくは請求項 2 記載の漢字変換装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は漢字変換装置に関し、とくに、入力文書に応じて基本辞書及び複数の分野別辞書を適切に選択する漢字変換装置に関する。

【0002】

【従来の技術】以下、本発明を従来の技術、実施例等により説明するに先立ち、重要な単語、用語等について説明、定義をしておく。本発明の主な対象は、中国語ワードプロセッサ、日本語ワードプロセッサ等の文字変換装置である。このため、本明細書にいう使用者若しく

は適用業務プログラム等により入力される漢字変換の対象となる「表音文字」とは、アルファベットのような純粋な表音文字の他、日本語における「かな」のごとき音素文字はもとより、アラビア数字のごとき表意文字的な文字をも包み、その他朝鮮語におけるオンモン（ハングル文字）、中国語における字音記号、例えば音調記号「一」、「／」、「V」、「\」に類似した記号（これらは、電子情報処理システムによる入力の都合で、類似した記号で代用している。）をも含む。また、漢字変換の結果たる「漢字」とは「法」、「特許法」等の如く純粋な「漢字」、より正確には「表意文字」や漢字のみから構成される単語のみならず、日本語における「愛する」等漢字と仮名とが混じった単語、句、節音をも包むのは勿論、その他「々（のま）」等の記号をも含む。

【0003】次に、「表音文字列」等の如く「列」という漢字を使用することがあるが、これは実際問題としてたとえ一の漢字でも複数の表音文字にて変換対象として指定されることが多いこと、変換結果も複数の漢字であることが多いことを考慮したことによるものであり、特に複数と限定しているわけではない。すなわち、一個の場合をも含む。また「語」、「単語」等も唯一の文字より構成されることもありえる。

【0004】更に「分野別辞書」における「分野」とは、「相互に関連のある単語、句、等を一まとめにした」とでもいうべき概念であり、「仏教」、「共産主義」、「法律」、「軍事」、「外交」等学問や実務における分類のみならず、「明代」、「漢代」、「周以前」等時代における区分、「北京語」、「上海語」、「日本語」、「朝鮮語」等言語における区分け、その他「個人名」、「芸能人」、「中国語化した外国語」等特殊な区分け等も含む。

【0005】漢字の読みによる入力方式は例えば、キーボードを通じて漢字の読みを表わす表音文字を入力し、辞書を参照して、形態素の読みコードの切り出しと文字変換処理を行ない、入力された表音文字列を文字列に変換する。しかるに、辞書に多くの字、語が含まれている場合には、形態素の読みコードの切り出し間違いが生じやすいし、同音字、語が増加するのと辞書の検索速度が低下することのため、漢字変換の効率も悪くなる。このため一般には、辞書を日常的、一般的な単語等を収納（集めた上、利用可能なように記憶する）した基本辞書と各分野で使用される単語を収納した分野別辞書に分けることがなされる。この場合には、使用者は入力したい文書の分野に応じて分野別辞書を選択することとなる。ところが、入力したい文書がいくつかの分野にまたがる場合には、使用者は絶えず分野別辞書を切り変える必要があるため、入力効率が悪くなる。さて、従来、この分野別辞書の切り換え機能を備えた漢字変換装置としては、例えば特開平 4-80863 号公報に示されたようなものがあつた。図 5 はこのような漢字変換装置の一例

10

20

30

40

50

の構成図である。本図において、1は入力部であり、2は制御部であり、3は表示部であり、4はテキストバッファであり、5は文字変換部であり、6は辞書検索部であり、70は基本辞書であり、71～7nは分野別辞書であり、8は辞書選択順序指定部であり、91及び92は辞書群保持部である。辞書群保持部、91、92は基本辞書70及び複数の分野別辞書71～7nを使用者が作成する文書の内容に応じた検索順序に従って保持する。入力部1は、仮名文字列などを入力したり、仮名漢字変換時に辞書群保持部91、92を切り換える操作を行なうキーボードを有する。テキストバッファ4は、入力された表音文字列や変換した漢字列を保持する。表示部3は、CRT等を有し、テキストバッファ4の内容などを表示する。辞書選択、順序指定部8は、基本辞書70及び複数の分野別辞書71～7nの中から作成する文書の内容に応じて検索したい複数の辞書の組をその検索順序に従って選択し、辞書群保持部91、92に登録する処理を行なう。辞書検索部6は、入力された仮名文字列を仮名漢字混じり文字列に変換するために辞書群保持部91、92に保持されている辞書を検索する。仮名漢字変換制御部5は、辞書検索部6で検索された辞書に基づいて入力された仮名文字列を仮名漢字混じり文字列に変換する。制御部2は、入力部1、テキストバッファ4、表示部3、辞書選択、順序指定部8、仮名漢字変換制御部5等を制御する。

【0006】以上のように構成された分野別辞書の切り換え機能を備えた漢字変換装置について、その動作を以下に説明する。使用者は、辞書選択、順序指定部8によって基本辞書70及び複数の分野別辞書70～7nの中から、作成する文書の内容に応じて検索対象となる漢字、単語が登録されている可能性の高い複数の辞書を選択し、更にその検索順序をも指定し、その結果を辞書群保持部91、92に登録する。例えば、情報処理に関する文書を作成する場合には、辞書群保持部91に通常の変換時に検索したい辞書（基本辞書と例えば文書の内容に応じての情報処理用語辞書）を、辞書群保持部92に固有名詞等に変換したいときに検索したい辞書（例えば、姓名辞書、地名辞書、基本辞書）を、各々この優先順序で登録しておく。入力部1から入力された仮名文字列は、テキストバッファ4に保持され、表示部3に表示される。仮名漢字変換処理においては、仮名漢字変換制御部5は、辞書検索部6によって辞書群保持部91、92に登録されている辞書を検索しながら、仮名文字列を仮名漢字混じりの文字列に変換する。そして、テキストバッファ4内の仮名文字列を変換後の文字列で置き換え、表示部3に表示することとなる。この際、上記の辞書を辞書群保持部91、92に登録しておいた場合には、通常の仮名漢字変換時には、辞書群保持部91に登録されている辞書が登録順に検索されるため、情報処理用語が一般用語よりもさきに出力される。もしも文書中

に人名などの固有名詞が出てきたならば、入力部1の辞書群切り換えキーを操作して検索される辞書群を辞書群保持部92に登録されている辞書に切り換える。もう一度辞書の切り換えキーを操作すれば辞書群保持部91に戻る。以上のようにして、辞書群切り換えキーを操作することにより、辞書群の切り換えを行なっている。

【0007】なお、同一発音の漢字列、語が複数ある場合に、それらを優先度の順に若しくは一度に表示部に表示した上で使用者に意図するものを特定させるための手順及びそのために必要な構成、使用者が特定した漢字列を目下作成中の文章の所定の位置に表示するための手順及び構成、入力された多数の表音文字からなる表音文字列を一体とした漢字変換の対象として区分けするために必要な手順、例えば最長一致法やそのための構成は、通常の中国語ワードプロセッサや日本語ワードプロセッサと同じ若しくは基本的には同じであり、また本発明の主旨には直接は関係しない。このため、それらについての図示や説明は省略する。

【0008】

【発明が解決しようとする課題】しかしながら、上記の従来技術に係る装置においては、以下のような問題点がある。第1の問題点。人手によって入力文書の内容に応じて辞書を選択し、且つ、分野別辞書を切り換える必要がある。しかしながら、分野別辞書の正しい切り換えは、使用者にとって大きな負担となる。また、辞書群保持部に登録されている辞書の分野が正しくない場合には、最長一致法等による漢字変換の対象となる形態素の読みコード（表音文字列）の切り出し及び文字変換が不正確となり、漢字変換効率も悪くなる。

【0009】第2の問題点。もし、入力文書にいくつもの分野の単語が含まれており、そのため分野の選定を煩雑に変更する必要がある場合には、使用者にとって分野別辞書の切り換えが大きな手間となる。本発明は、上記の問題点に鑑み、適切かつ自動的に分野別辞書の選定がなされる漢字変換装置を提供することを目的としてなされたものである。

【0010】

【課題を解決するための手段】上記目的を達成するために、請求項1の発明については、入力された表音文字列に相応する形態素の読みコードを切り出して辞書検索部へ送る入力部と基本辞書と複数の分野別辞書とからなる辞書部と表示部とを有する漢字変換装置において、各辞書についてその使用回数を記録する辞書使用記録部と、上記辞書使用記録部に記録されている使用回数を多いものを優先して各辞書の検索順序を決める辞書選択部と、上記入力部により送られてきた読み記号を検索キーとして、上記辞書選択部が決める検索順序に従って各辞書の検索を行ない、読み記号に対応する全ての同音字、語候補及び当該同音字、語候補毎にそれらが登録されている分野別辞書のコードを取り出して文字変換部へ送る辞書

検索部と、上記辞書検索部により検索された同音字、語の候補を表示部に表示させ使用者に正しい変換を確認させると共に、正しいとされた同音字、語の登録されている辞書のコードを辞書管理部へ送る文字変換部と、上記文字変換部から送られてきた辞書コードをもとに、上記辞書使用記録部の記録する各辞書の使用回数を所定の手順で更新する辞書管理部とを有していることを特徴とする漢字変換装置としている。

【0011】請求項2の発明においては、前記辞書管理部は、上記辞書使用記録部の記録を更新する際に、使用回数が記録されている複数の辞書のうち、所定数のものの使用回数があらかじめ定められたしきい値に達したときには、それまでの若しくは一まとまりの漢字変換の終了したときまでの使用回数がそのしきい値に達していない辞書に対してはその使用回数から一定値を引き去るか0にし、それまでの使用回数が0か、前記引き去りの結果負となった辞書に対しては不使用に設定することを特徴とする請求項1記載の漢字変換装置としている。

【0012】請求項3の発明においては、前記表示部は、前記辞書検索部が所定数の辞書で読み記号に対応する同音字、語候補の検索が成功した場合には、その段階でその検索に成功した同音字、語候補のみを使用者による正しい変換の確認のために表示部に表示することとを特徴とする請求項1若しくは請求項2記載の漢字変換装置としている。

【0013】

【作用】上記構成により、請求項1の発明においては、辞書使用記録部は各辞書毎にその使用回数を記録（使用状況に応じて変化する形での記憶）する。辞書選択部は、上記辞書使用記録部に記録されている各辞書の使用回数の多いものを優先して、その検索順序を決める。辞書検索部は、入力部より送られてきた読み記号を検索キーとして辞書選択部が決めた検索順序に従って各辞書の検索を行ない、読み記号に対応して収納されている全ての同音字、語候補（この段階では、使用者の意図する字、語の候補である）及び当該同音字、語候補毎にそれらが登録（半永久的若しくは永久的な形であらかじめの記憶）されている分野別辞書のコードを取り出して文字変換部へ送る。文字変換部は、辞書検索部から検索された同音字、語候補を表示部に表示させ、使用者に正しい変換を確認させると共に、正しいとされた同音字、語の登録されている辞書のコードを辞書管理部へ送る。辞書管理部は、文字変換部から送られてきた辞書コードをもとに、所定の手順で辞書使用記録部の記録する各辞書の使用回数を更新する。

【0014】請求項2の発明においては、辞書管理部は、上記辞書使用記録部の記録を更新する際に、使用記録されている複数の辞書のうち、所定数のものの使用回数があらかじめ定められたしきい値に達したときには、それまでの若しくは一まとまりの漢字変換の終了したと

きまでの使用回数がその定められた値に達していない辞書に対してはその使用回数から一定値を引き去り、その結果それまでの使用回数が0か、負となった辞書に対しては不使用に設定する。

【0015】請求項3の発明においては、表示部は、辞書検索部が所定数の辞書で読み記号に対応する同音字、語候補の検索が成功した場合には、他の辞書の検索の実施の有無にかかわらず、その段階でその検索に成功した同音字、語候補のみを使用者による正しい変換の確認のために表示部に表示する。

【0016】

【実施例】以下、本発明を実施例に基づき説明する。

（第1実施例）本実施例は、請求項1及び請求項2の発明に関する。図1は、本発明に係る漢字変換装置の一実施例の構成図である。本図において、1は入力部であり、2は制御部であり、3は表示部であり、4はテキストバッファであり、5は文字変換部であり、6は辞書検索部であり、70は基本辞書であり、71～7nは分野別辞書であり、81は辞書管理部であり、82は辞書使用記録部であり、90は辞書選択部である。制御部2、表示部3、テキストバッファ4、基本辞書70及び分野別辞書71～7nの構成、原理は従来技術に係わるものと同じ若しくは大差がないため、これらについての説明は本発明に関係しない限り省略する若しくは簡単なものにとどめる。基本辞書70は、読み記号を検索キーとしてこれに対応する基本的な語を収納している。辞書使用記録部81は、辞書コードを検索キーとして使用回数の値若しくは不使用記号“X”の値を記録し、且つ、その使用回数の値が大のものを小のものより先の順に記録する。分野別辞書71～7nは、複数の分野に対して、各分野固有の単語を収納しているものであり、読み記号を検索キーとしてその分野で対応する単語を収納している。辞書選択部90は、辞書使用記録部82に記録されている各分野別辞書の使用回数及び配列順序を参照して、使用する分野別辞書及びその検索順序を決め、また辞書管理部81の指示により、検索対象とする分野別辞書と検索順序を決める。辞書検索部6は、読み記号を検索キーとして、辞書選択部90で指定された分野別辞書及びその検索順序に従って対応する全ての同音字、語候補を取り出す。辞書管理部81は、使用者により正しい変換であると確定された文字列における漢字、単語毎に、それらの辞書コードをもとに辞書使用記録部82に記録されている各辞書の使用回数の記録及び配列の順序を更新処理する。文字変換部5は、入力された表音文字列を辞書検索部6を通じて変換候補の漢字列、単語に変換した上で表示部3に表示させ、この上で使用者に意図する漢字列、単語を選択させると共に、正しいとされた漢字列、単語の登録されている辞書のコードを辞書管理部81に送る。表示部3は、辞書検索部6の検索した同音字、語候補を順に表示し、更に使用者により正しい

とされた漢字列、単語を作成中の文章の所定位置に表示する。

【0017】図2及び図3は、本実施例の動作流れ図である。以下、図2をもとに、その基本的な動作流れを説明する。

(S21) 使用者が入力部より、一まとまりの漢字変換の対象としての表音文字列を入力する。なお、後に示す具体例では、7つの音節を構成する表音文字列が入力されている。

【0018】(S22) 入力部は、入力された表音文字列の中から、一体としての変換の対象となる単語に相応する形態素の読みコードを切り出す。なお、後に示す具体例では、2若しくは3の音節が単語を構成している。

(S23) 辞書検索部は、辞書使用記録部に記録されている使用回数を参照して、分野別辞書若しくは基本辞書内を検索し、これに対応する字、語候補を表示する。この場合、もし、未変換或いは変換の間違ったとされた形態素の全ての読みコードの検索であるならば、当該形態素に対して文字変換を行なうことになる。図4にその辞書使用記録部の使用回数の記録内容を概念的に示す。

【0019】(S24) 使用者は、表示部を参照して変換された漢字が全て正確か否かを判断する。もし、不正確なものがあればそれを指摘し、当該形態素やそれを含む一体として変換対象となる形態素の読みコードを参照して再度検索を行うべく、(S23)へ戻る。なければ、(S25)へ行く。

(S25) 形態素の読みコードに正しく対応するとされた漢字、単語を含む辞書の辞書コードを参照して辞書使用記録部の内容を更新する。その更新の際の詳しい動作流れについては、後で詳しく説明する。

【0020】(S26) 正しく変換された漢字列を作成中の文章の所定位置に表示する。次に、図3をもとに辞書管理部の動作流れを説明する。

(S301) しきい値THと一まとまりの漢字変換の対象として入力された表音文字列の中でその中に含まれている一体としての漢字変換となった単語の総数hを設定する。

【0021】(S302) 正しい変換であると確定された単語に相応する漢字、単語の辞書コードを順に一時記憶装置に格納する。このため、変数A[J]が設定される。ここに、J=1, 2, ..., hであり、各A[J]の内容は辞書コードである。

(S303) 変数Jの開始値を1に設定する。

(S304) J>hか否かを判断する。もし、J>hでなければ、(S305)へ行く。

【0022】(S305) A[J]の値を検索キーとして辞書使用記録部を参照して該当する辞書の使用回数を取り出し、変数としてCNTを記憶する。なお、CNT=Xであれば、当該辞書は不使用であるため、かかる場合はありえない。

(S306) CNTに1を足すと上記(S307)へ行く。

【0023】(S307) CNT>THか否かを判断する。もし、CNT>THであれば更新処理を行うこととなるため、(S308)へ行き、CNT>THでなければ(S310)へ行く。

(S308) 現在辞書使用記録部に記録されている辞書の使用回数から一定値Dを引き去る。この上で、(S309)へ行く。

【0024】(S309) その結果を新しく使用回数に設定する。使用回数が0若しくは負である辞書は、その使用回数をXに設定する。(S310)へ行く。

(S310) A[J]を検索キーとして該当する辞書をもとめ、この新しいCNTの値を辞書使用記録部に記録する。

(S311) Jに1を足した上で、次の漢字、単語を対象として同じ処理を行うべく(S304)へ戻る。

【0025】(S312) 使用回数をもとに辞書使用記録部の使用回数の記録と配列の順序の更新が終了する。

図4は、辞書使用記憶部の記憶内容を概念的に示す図である。本図において、第1番目の列は分野別の辞書コードを示すものであり、第2番目の列は使用回数を示すものであり、これらは上から下の行へ使用回数の大きい順に、配置されている。図中、使用回数がXであるものは当該分野別辞書が不使用とされていることを示す。

【0026】なお、本実施例では、基本辞書は本発明に係わる作用の対象外かつ第一の優先順位で検索され、辞書管理部のしきい値と使用回数の比較がなされる分野別辞書の所定数は1とされている。実際問題としていずれの分野別辞書よりも基本辞書の使用頻度が多いため、これは基本辞書をも本発明に係る作用の対象とした上で、しきい値と使用回数の比較がなされる辞書の所定数を2としたものと実質同じである。

【0027】ただし、この場合には基本辞書の使用頻度のみが多くなるため、その使用回数の記録の数字がどうしても多くなりがちであり、このためその上限を設ける等その処理対策が必要となる。次に、図4に示す辞書使用記憶部を例にとって、図3に示す辞書管理部の動作流れにより、これがどのように変化するかを具体的に説明する。

【0028】今、「lian2jyuen1zuen1shou3guo2ji4fa3」という表音文字列が入力されると、第1に構成する音節の数を、第2に入力順位を優先させて漢字変換の対象としての音節を切り出す最長一致法により形態素の読みコードの切り出しを行なった結果は「lian2jyuen1 zuen1shou3 guo2ji4fa3」となる。そして、分野別辞書と基本辞書を参照して、上記表音文字列を「聯軍遵守國際法」という文字列に変換する。その関連する辞書コードはそれぞれ「聯軍」が5(軍事)で、「遵守」が0(基本辞書)で、「國際法」が7(法律)である。なお、先

の図 3 に示す動作流れの説明において、各 A[J] の内容は「分野別辞書のコード」のみとなり、分野別辞書の種類は全部で 9 個であるため、(S 3 0 2) における A[J] の値は 1 から 9 までとなる。ここで、辞書コードが「0」であるものは基本辞書を示す。

【0029】次に、図 3 の辞書管理部の処理流れの説明に入る。

(S 3 0 1) 上記一連の表音文字列の変遷で一体として漢字変換の対象となった漢字列、単語の個数は 3 であるため $h = 3$ とされる。易しく説明するために、こ

こで、TH の値を 137 に設定する。

(S 3 0 2) A[1]=5, A[2]=0, A[3]=7

(S 3 0 3) 変数 J を 1 に設定する。

【0030】(S 3 0 4) $J < 3 = h$ であるので、(S 3 0 5) に行く。

(S 3 0 5) A[1] の値 5 を検索キーとして辞書使用記録部からその使用回数 137 を取り出して CNT に記憶する。

(S 3 0 6) CNT に 1 を足し、138 となる。

(S 3 0 7) $CNT > TH$ のために、(S 3 0 9) に入る。

【0031】(S 3 0 8) 辞書使用記録部は一定値 10 を引き去ることにより以下のように更新される。(ただし、図 4 における更新部分のみを示す。そして、これは後の更新内容を示す場合も同様である。)

5 (軍事) 138

3 (経済) 59

7 (法律) 1

8 (会計) X

以上のもので、(S 3 1 0) へ行く。

【0032】(S 3 1 0) A[1] の値 5 を検索キーとして辞書をもとめ、現在の CNT の値を使用回数として辞書使用記録部に記録する。

(S 3 1 1) J に 1 を足し 2 としてから、上記 (S 3 0 4) へ戻る。J = 2, J = 3 においても同じ処理動作を行なう。そして、この状態での辞書使用記録部の内容を順に以下に示す。

【0033】J = 2 :

5 (軍事) 139

3 (経済) 59

7 (法律) 1

8 (会計) X

J = 3 :

5 (軍事) 139

3 (経済) 59

7 (法律) 2

8 (会計) X

この場合、(S 3 0 4) では J = 4 であり、 $4 > 4$ のために、(S 3 1 2) に行く。

【0034】(S 3 1 2) では、使用回数の記録とそれ

に応じての配列順序の変更の処理を終了する。この場合、最終的に使用されている分野別辞書及びその配列順序は、同じである。

(第 2 実施例) 本実施例は、先の第 1 実施例において、辞書管理部が、ある分野別辞書の使用回数があらかじめ定められたしきい値に達したときには、それまでの使用回数を 0 に設定し直す他は第 1 実施例と同じである。

【0035】このため、各部の構成図を図示した上でその動作作用等を説明することや、処理流れ図における同じ動作の説明は省略し、本実施例固有の動作、処理流れ等のみ説明する。先の (S 3 0 7) において、辞書管理部が全ての分野別辞書の値を 0 にする。このため、辞書使用記録部の内容は以下になる。

【0036】5 (軍事) 0

3 (経済) 0

7 (法律) 0

8 (会計) 1

同じく、(S 3 1 1) において、J = 2 の動作が終了した段階では、辞書使用記録部の内容は、以下になる。

【0037】5 (軍事) 1

3 (経済) 0

7 (法律) 0

8 (会計) X

J = 3 の動作が終了した段階では、同じく以下になる。

【0038】5 (軍事) 1

3 (経済) 0

7 (法律) 1

8 (会計) X

以上の使用回数のもとで、辞書選択部は各分野別辞書の配列順序を変更する。このため、最終的には、以下になる。

【0039】5 (軍事) 1

7 (経済) 1

3 (法律) 0

8 (会計) X

先の第 1 実施例と比較した場合に、本実施例では全ての分野別辞書の使用回数を 0 に設定し直すため、当初から使用するものとされた分野別辞書の設定数 (種類) が多い場合に有利となる。

(第 3 実施例) 本実施例は、請求項 3 の発明に関する。

【0040】本実施例は、基本的構成は先の第 1 実施例と同じである。このため、構成図等は省略する。ただ、辞書選択部が計数部を内蔵し、この計数部は検索キーたる読み記号毎に検索対象とされる辞書の個数を計算し、この値があらかじめ定められた一定値内で検索に成功したならば、この旨を文字変換部に通知するようになっている。また、この通知を受けた文字変換部は、その段階で検索された同音字、語候補を使用者による正しい選択

の確認のため表示部の所定位置に表示させる。またこのため、それ以後優先順位の低い辞書を検索することにより検出された同音語、語候補を表示部に表示せず保持しておくための一時記憶部を有している点が異なる。

【0041】ただし、計数部、計数部からの通知により文字変換部を途中の段階で作用発揮させるための引き金となる信号（トリガー信号）発生部、一時記憶部の構成は、いわば周知の技術であるため、その内容の説明は省略する。先の実施例と比較した場合に、本実施例では、検索の途中段階で正確率の高い同音字、語候補のみを使用者の意図する漢字、単語の確認の対象として表示し、選択することがなされるため、それだけ漢字変換の速度が向上する。

【0042】以上、本発明を実施例に基づき説明してきたが、本発明は何も上記実施例に限定されないのは勿論である。すなわち、例えば以下のようなものも本発明に包含され、

（1）第1、第2実施例では、使用しない分野別辞書の表示にXを付すこととしたが、その他の任意のフラグ、手段を採用する。同じく、各分野別辞書の使用回数からの一定値の引き去りは、各単語毎に（具体例におけるA[J]、J=1、2、3毎に）なすのではなく、一まとまりの漢字変換の対象とされた全ての単語の各分野別辞書についてその使用回数を追加訂正した上で、（具体例における「聯軍」、「遵守」、「国際法」全てについて訂正した上で）しきい値と順に比較していき、一の分野別辞書でもその使用回数がしきい値に達したならば、一定値の引き去りを行うようにしている。

（2）制御レジスタを使って各辞書毎の更新回数を累加することにより、その値があるしきい値より大であるときに辞書使用記録部を更新する。

（3）各請求項の発明において、基本辞書は辞書使用記録部、辞書選択部、辞書管理部、辞書検索部の作用の対象外とする若しくは使用者の意思により作用対象外とすることやどの分野別辞書を基本辞書と指定することが可能である機能が付加されている。

【0043】同じく、使用者が特定の分野別辞書については最初から作用対象外とすることが可能な機能が付加されている。その他、請求項2の発明における所定数、しきい値、一定値を変更可能な機能が付加されている。

（4）分野別辞書は、装着、取りはずしが可能な機能が付加されている。

（5）辞書検索部が検索した同音字等を文字変換部が表示部に表示させる際に、多数の字等が検索されたためそのままではこれらすべてを表示させるのが困難な場合には使用頻度の高いものから表示していく、使用頻度の高いものを大きく表示し、低いものを小さく表示することにより全てを一度に表示する等の機能が付加されている。

（6）基本辞書、各分野別辞書毎に学習能力を発揮しえ

る装置が付加されている。

（7）辞書使用記録部の記録内容そのものを別途の装置に記録の上必要に応じて呼び出し、切り換えを可能とする。これにより、作成する文書の分野如何によっては以前作成した同一分野の文書の使用経験を現在作成中の文書に反映可能、かつ同一のワードプロセッサ等で、多分野の文書を効率よく作成可能とする。

（8）製造等の都合で、特許請求の範囲に記載した一の構成要素（要件）を物理的、機械的に複数の物としている、逆に複数の構成要素を一の物としている、あるいはこれらを適宜組み合わせている。

（9）同音字、語候補が検出されなかった場合には、この旨を使用者に表示すると共に自動的に不使用とされている分野別辞書を検索する機能発揮のための選択機構が付加されている。

【0044】

【発明の効果】以上説明してきたように、本発明によれば、以下のような効果が発揮される。

（1）人手によって分野別辞書を選択する必要がなくなり、多くの分野の単語、漢字を含む文書の作成に対して漢字変換の効率を高める。

（2）分野別辞書を対象としての使用経験に応じた学習能力が得られ、必要な分野別辞書を優先して選択し、不必要な辞書を閉じることにより、辞書の検索速度が向上する。

【0045】（3）常に、必要な分野別辞書のみを保持することにより、漢字変換の候補として検索し、また列挙する同音漢字、語の候補も少なくなる。このため、文字変換の正確率と速度を大きく高めるだけでなく、使用者の目等の疲労もすくなるなる。

【図面の簡単な説明】

【図1】本発明に係る漢字変換装置の一実施例の構成図である。

【図2】上記実施例の処理流れ図である。

【図3】上記実施例の辞書管理部の動作流れ図である。

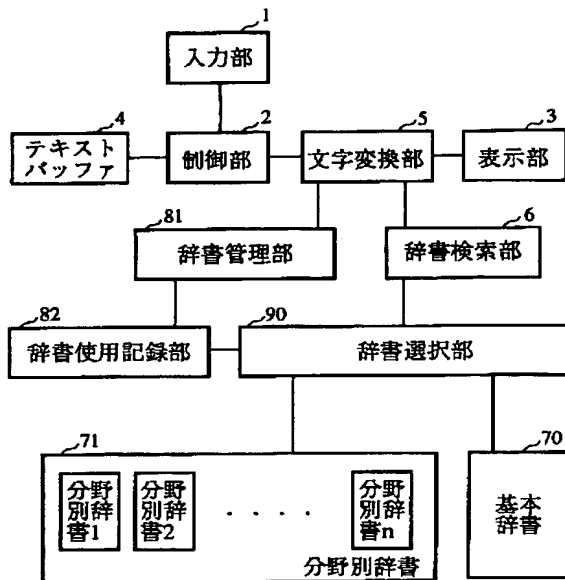
【図4】上記実施例の辞書使用記録部の内容を概念的に示す図である。

【図5】従来の漢字変換装置の構成図である。

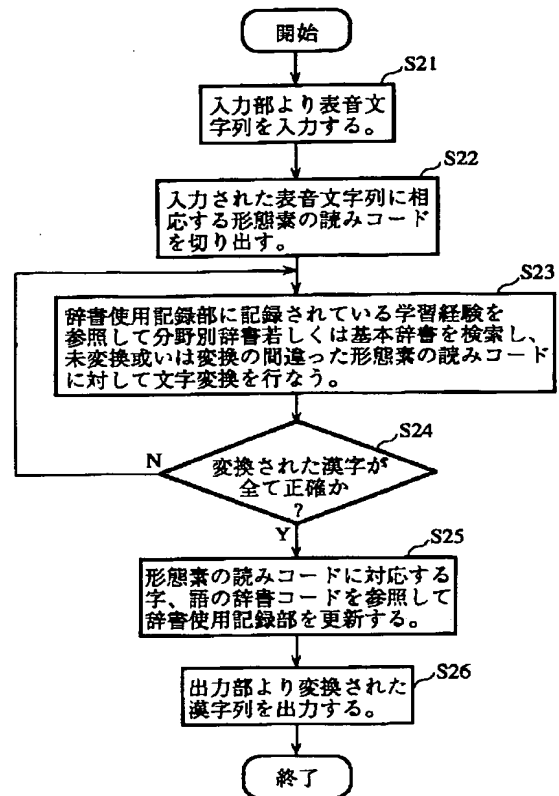
【符号の説明】

- 1 入力部
- 2 制御部
- 3 表示部
- 4 テキストバッファ
- 5 文字変換部
- 6 辞書検索部
- 7 0 基本辞書
- 7 1、・・・、7 n 分野別辞書
- 8 1 辞書管理部
- 8 2 辞書使用記録部
- 9 0 辞書選択部

【図 1】



【図 2】



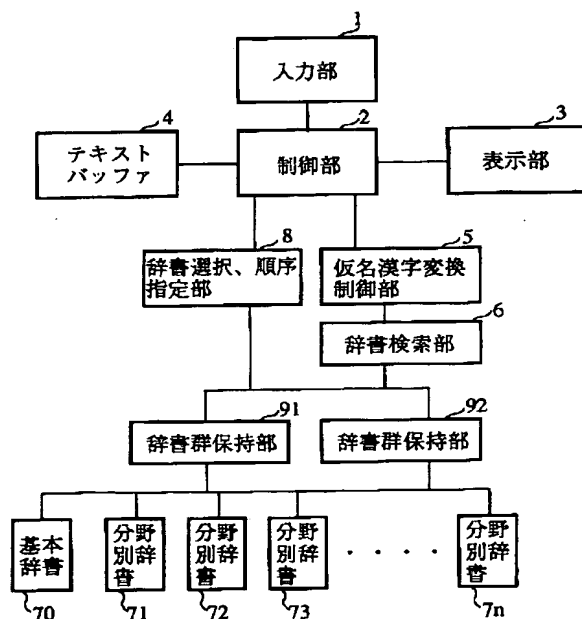
【図 4】

分野別辞書コード 使用回数

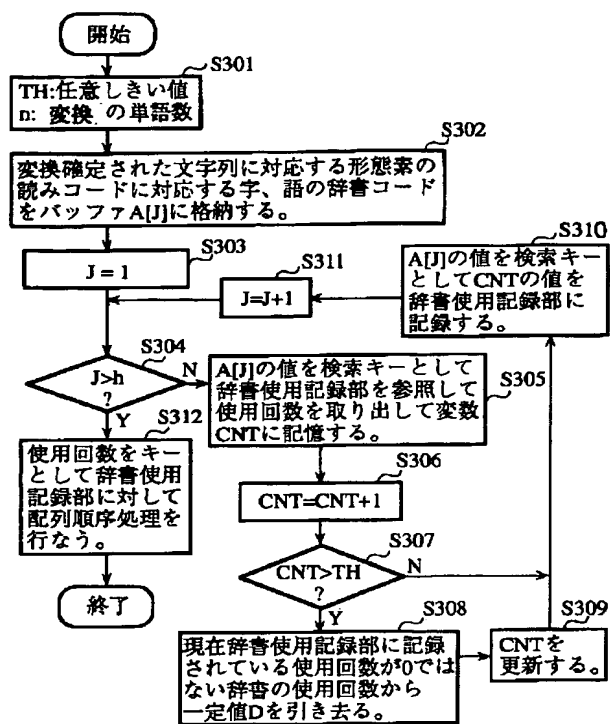
5 (軍事)	137
3 (経済)	69
7 (法律)	11
8 (会計)	1
1 (教育)	X
9 (医学)	X
4 (機械)	X
2 (政治)	X
6 (生物)	X

(X: 不使用)

【図 5】



【図 3】



加修正其使用次数后(具体例中,对“联军”、“遵守”、“国际法”全进行修正后),依次比较阈值,只要-专业辞典其使用次数达阈值,就进行减去规定值。

(2)通过使用控制寄存器累加各辞典每次更新次数,当其值大于某阈值时也更新辞典使用记录部。

(3)在本发明各形态中,基本辞典除作为辞典使用记录部、辞典选择部、辞典管理部、辞典检索部的作用对象外或者除按使用者意思成为作用对象外还能添加将某专业辞典指定为基本辞典的功能。

同样,使用者对特定的专业辞典除从起初就成为作用对象以外,还能添加其它可能的功能。此外,也可增加可变更本发明第二形态的规定数、阈值、规定值的功能。

(4)添加能装、卸专业辞典的功能。

(5)字符变换部让辞典检索部检索到的同音字等显示于显示部时,添加在由于检索到许多同音字而难以直接全部显示的情况下从使用频度高的开始进行显示的功能,或通过让使用频度高的显示得大,低的显示得小而一次性进行显示等功能。

(6)对基本辞典、各专业辞典分别添加可发挥学习能力的装置。

(7)将辞典使用记录部的记录内容本身记录在其他途径的装置上并可根据需要而读出、转换。因此,能根据写作文件的专业情况,将以前写过的高一专业文件的使用经验反映至现在写作中的文件中,而且能在同一文字处理机等上有效作成多专业的文件。

(8)为便于制造,可将本发明所述形态中的一构件要素(构成要件)在物理和机械方面分为多个构件,反之也可将多个构件做成一个构件,或者将它们适当组合。

(9)在检不出同音字、候选词时,添加选择机构,以便将此情况向使用者显示并自动发挥作用检索被定为不使用的专业辞典。

综上所述,若利用本发明,则具有以下效果。

(1)不必由人工选择专业辞典,对于包含许多专业的单词、汉字

的文件写作, 汉字变换效率高。

(2) 由于能获得以专业辞典为对象的使用经验方面的学习能力, 通过优先选择必要的专业辞典、关闭不必要辞典, 提高辞典的检索速度。

(3) 由于平时只保持必要的专业辞典, 作为汉字变换的候选而检索及列举的同音汉字、候选词也变少。因此不仅大为提高字符变换正确率和速度, 而且还减轻使用者的用眼疲劳等。

说明书附图

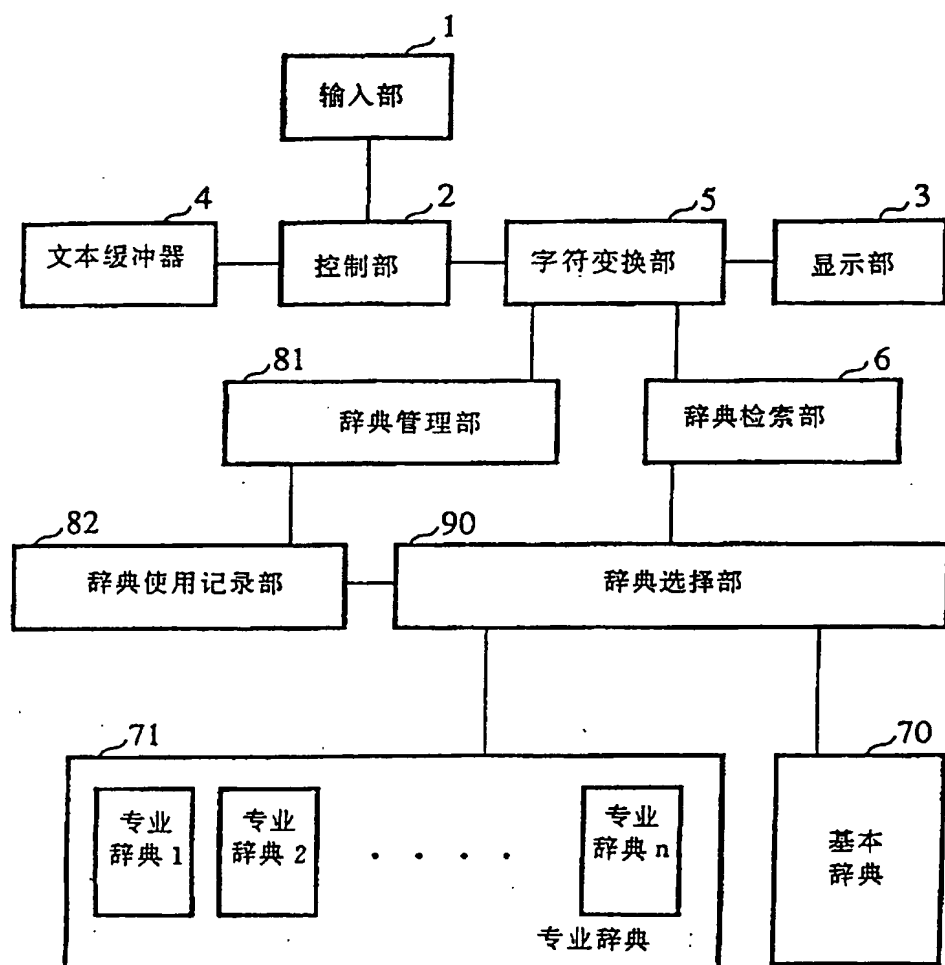


图 1

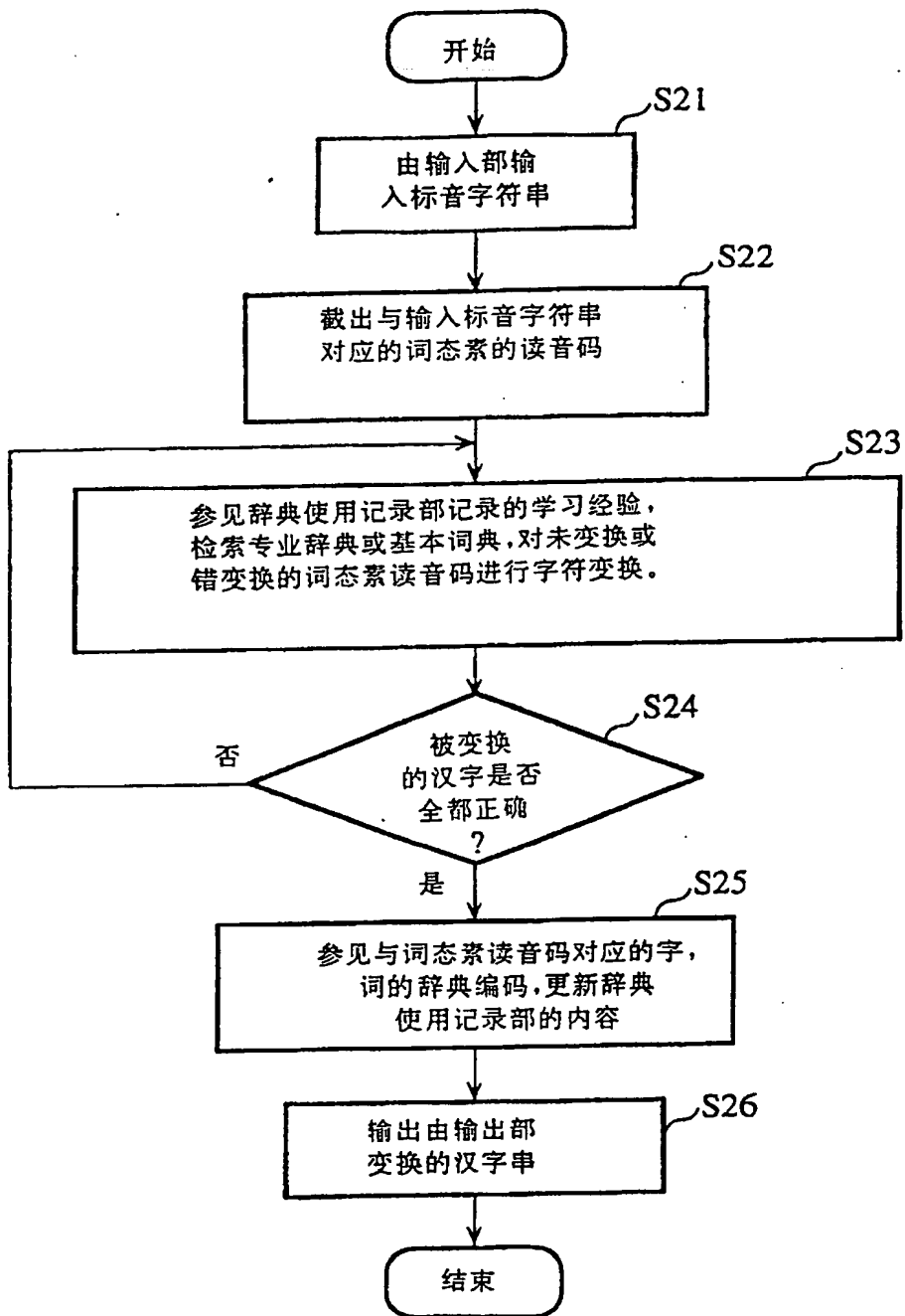


图 2

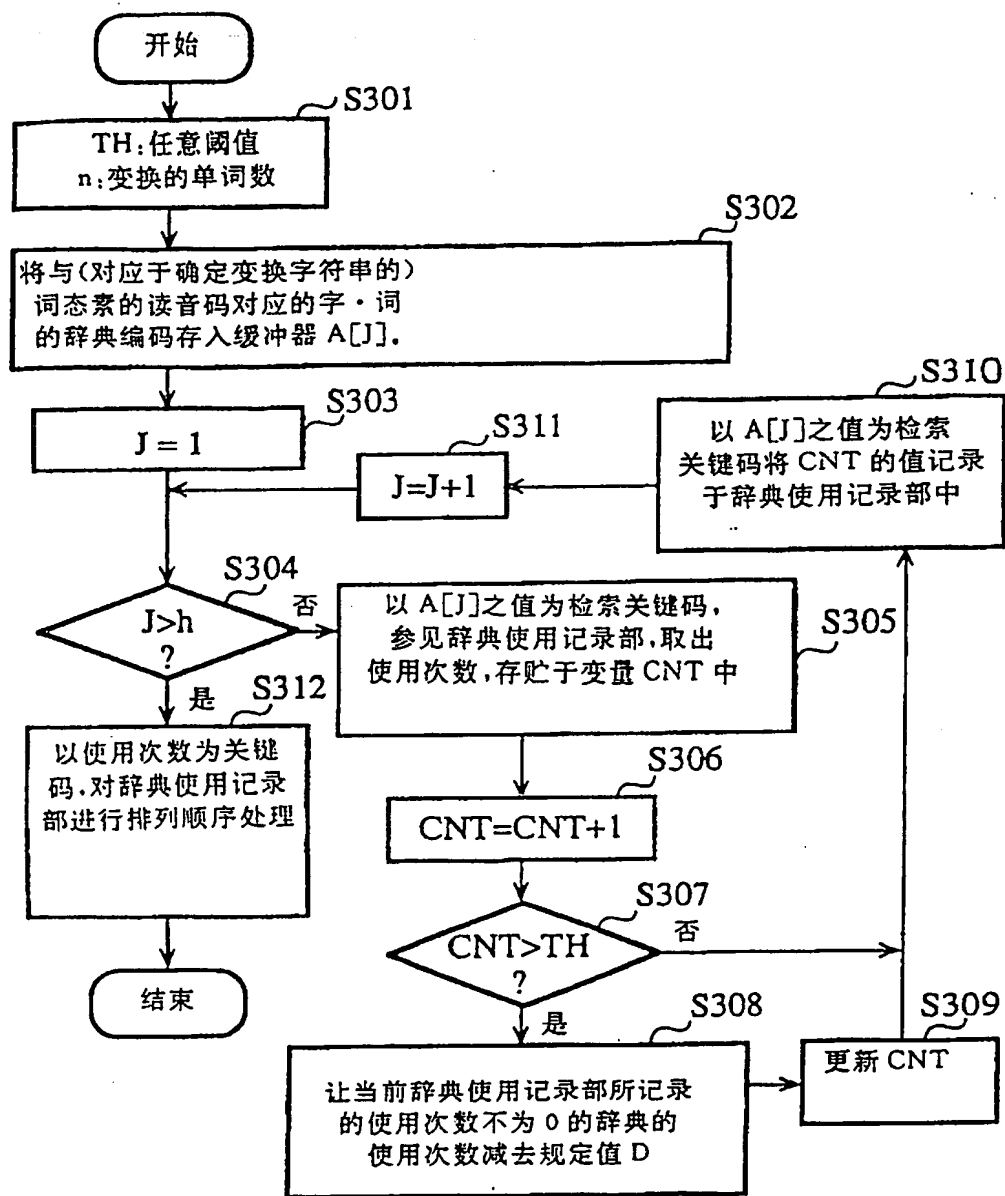
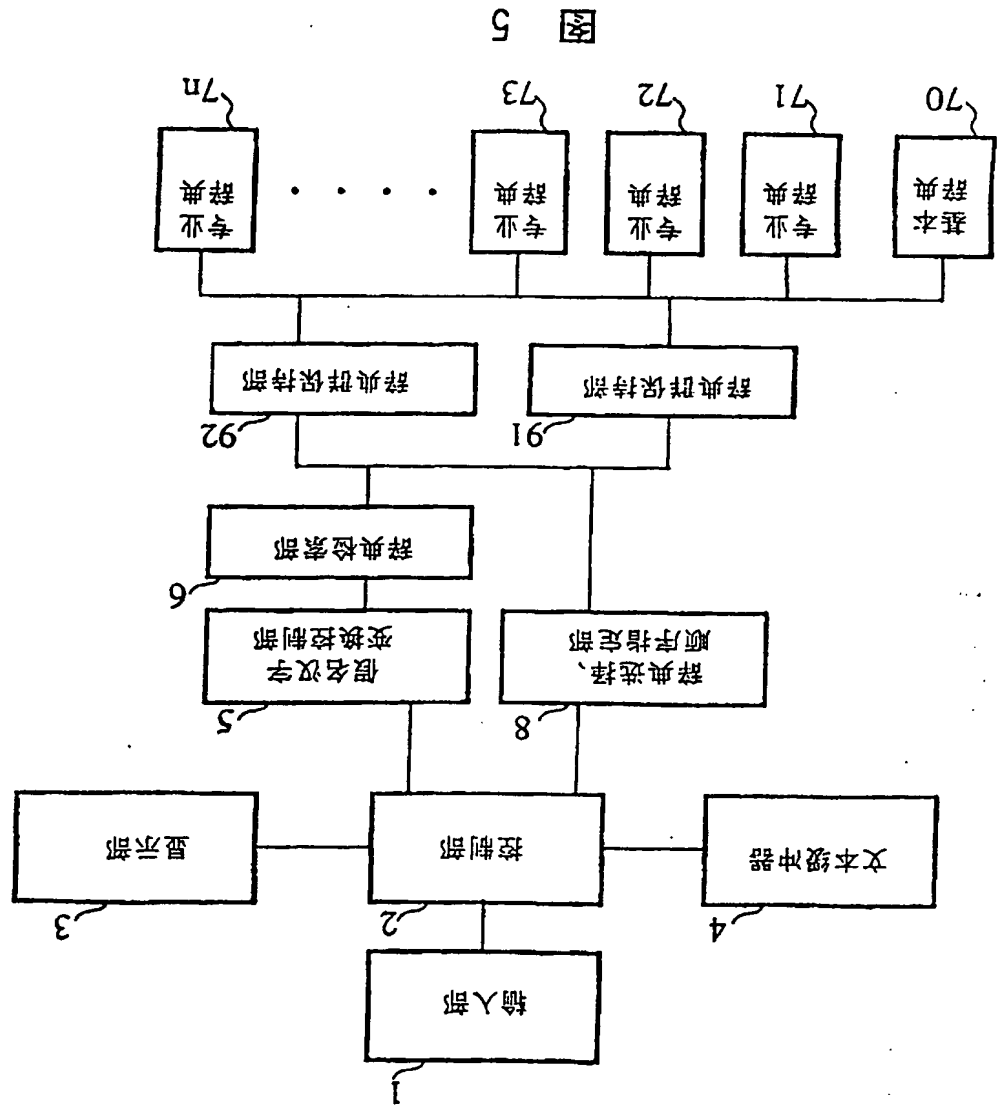


图 3

专业辞典编码	使用次数
5 (军事)	137
3 (经济)	69
7 (法律)	11
8 (会计)	1
1 (教育)	X
9 (医学)	X
4 (机械)	X
2 (政治)	X
6 (生物)	X

(X: 不使用)

图 4



[12]发 明 专 利 说 明 书

[21] ZL 专利号 94104792. X

[45]授权公告日 2002 年 5 月 8 日 [11]授权公告号 CN 1084500C

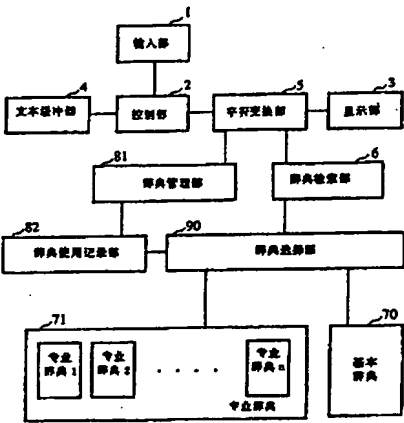
[22]申请日 1994.5.7 [24]颁证日 2002.5.8
[21]申请号 94104792. X
[30]优先权
[32]1993.11.2 [33]JP [31]274604/93
[73]专利权人 松下电器产业株式会社
地址 日本大阪府
[72]发明人 郭俊桔
[56]参考文献
EP 0366142A2 1990. 5. 2 G06F15/38
EP 0370774A2 1990. 5.30 G06F15/38
审查员 李 琼

[74]专利代理机构 上海专利商标事务所
代理人 汪 瑜

权利要求书 1 页 说明书 13 页 附图页数 5 页

[54]发明名称 汉字变换装置
[57]摘要

汉字变换装置。其输入部从输入标音字符串截出相应词态素读音码送至词典检索部。词典使用记录部分别对各专业词典记录使用次数。词典选择部让使用次数多的专业词典优先检索。词典检索部以截出读音码为检索码按优先顺序检出相应的所有同音字、候选词。由文字变换部将它们显示于显示部,供使用者确认正确的变换并把登记有认为正确的字、词的词典的编码送至词典管理部。词典管理部据此更新词典使用记录部。从而能恰当选择词典。



ISSN 1008-4274

权 利 要 求 书

1. 一种汉字变换装置具有将与输入标音字符串对应的词态素读音码截出后送至辞典检索部的输入部、由基本辞典与多本专业辞典构成的辞典部和显示部,其特征是还具有:①对各专业辞典记录其使用次数的辞典使用记录部;②让上述辞典使用记录部所记录的使用次数多的优先来决定各辞典检索顺序的辞典选择部;③以上述输入部送来的读音记号为检索关键码,按上述辞典选择部决定的检索顺序进行各辞典的检索,取出与读音记号对应的所有同音字、候选词以及登记有它们的专业辞典的编码,再送至字符变换部的辞典检索部;④使上述辞典检索部检出的同音字、候选词显示于显示部,让使用者确认正确的变换,同时将登记有被认为正确的同音字、词的辞典的编码送至辞典管理部的字符变换部;⑤根据上述字符变换部送来的辞典编码,按所定程序更新上述辞典使用记录部记录的各专业辞典的使用次数的辞典管理部。

2. 根据权利要求 1 所述的汉字变换装置,其特征是上述辞典管理部更新上述辞典使用记录部的记录时,在记录有使用次数的多本专业辞典中,专业辞典的使用次数达到预定阈值情况下,对迄今的或者汇总汉字变换结束前的使用次数未达到其阈值的专业辞典、将其使用次数减去规定值而为 0,或者迄今使用次数为 0 而经上述减法结果为负的专业辞典,设定为不使用。

3. 根据权利要求 1 或 2 所述的汉字变换装置,其特征是上述辞典检索部在辞典中检索到与读音记号对应的同音字、候选词时,上述显示部只将该阶段中检索到的同音字、候选词显示于显示部,以便使用者确认正确的变换。

说明书

汉字变换装置

本发明涉及汉字变换装置,特别是关于按照输入文件正确选择基本辞典和多个专业辞典。

为了用以往技术、实施例等说明本发明,先对重要的单词、用语等进行说明、定义。

本发明的主要对象是中文文字处理器、日文文字处理器等文字变换装置。因此,本发明说明书中所述的由使用者或应用业务程序等所输入的作为汉字变换对象的“标音字符”,除罗马字母那样的纯粹标音字符外,当然还包括像日文中假名那样的音素字符,另外还包括阿拉伯数字那样的表意文字的字符,还有朝鲜语中的谚文(朝鲜文字)、中文中的字音记号,例如类似音调记号“—”、“/”、“v”、“\”的记号(这些在电子信息处理系统的输入中以类似记号代用)。而且,作为汉字变换结果的“汉字”不仅包括如“法”、“特许法”的纯粹汉字、较确切地只由“表意文字”、汉字构成的单词,不用说还包括如“爱する”等汉字与假名混和的单词、句节,此外,还包括“々(のま)”等记号。

关于说明中常使用如“标音字符串”等词汇中出现的“串”这样的汉字,由于实际上常常是即使一个汉字也用多个标音字符指定变换对象,而且考虑到变换结果是多个汉字的情况为多,所以不特别限定为复数。即,也包含单个的情况。还有,“语”“词”等也往往由独一无二文字构成。

至于“专业辞典”中的“专业”的概念,可以说是“相互关连的单词、句子等的统称”,不仅包括“佛教”、“共产主义”、“法律”、“军事”“外交”等学术或者实际业务方面的分类,还包括“明代”、“汉代”、“周以前”等时代的划分、“北京话”、“上海话”、“日语”、“朝鲜语”等语言的区分、以及“姓名”、“艺人”、“中文化

的外语”等特殊区分。

汉字读音的输入方式,例如通过键盘输入表示汉字读音的标音字符,对照辞典进行词态素的读音码截出和文字变换处理,再将输入的标音字符串变换成文字串。然而,在词典中含许多字、语的情况下,词态素的读音码的截出容易搞错,会因同音字、语增加和辞典检索速度降低而降低汉字变换的效率。因此,一般将辞典分为收存(在收集基础上,存贮成可利用状态)日常的一般性单词的基本辞典和收存各专业所使用的单词的专业辞典。这种情况下,使用者就可以根据想要输入的文件的专业选择专业辞典。可是,当想输入的文件跨及几个专业时,使用者需经常转换专业辞典,因此输入效率变差。以往,具有这种转换专业辞典功能的汉字变换装置有日本特许公开号平 4—80863 号公开的装置。图 5 为这种汉字变换装置一例的结构图。该图中,1 为输入部、2 为控制部、3 为显示部、4 为文本缓冲器、5 为字符变换部、6 为辞典检索部、70 为基本辞典、71-7n 为专业辞典、8 为辞典选择顺序指定部、91 及 92 为辞典群保持部。辞典群保持部 91、92 按照与使用者编制的文件内容对应的检索顺序保持基本辞典 70 和多个专业辞典 71-7n。输入部 1 具有输入假名字符串等、或在假名汉字变换时进行切换辞典群保持部 91、92 的操作的键盘。文本缓冲器 4 保持输入标音字符串和变换后的汉字串等。显示部 3 具有 CRT 等,显示文本缓冲器 4 的内容等。辞典选择、顺序指定部 8 根据编制文件的内容并按照其检索顺序从基本辞典 70 及多个专业辞典 71-7n 中选择想检索的多个辞典组,进行登记在辞典群保持部 91、92 上的处理。辞典检索部 6 检索保持在辞典群保持部 91、92 中的辞典,以便将输入的假名字符串变换成假名汉字混和字符串。假名汉字变换控制部 5 根据辞典检索部 6 检索到的辞典将输入的假名字符串变换成假名汉字混和字符串。控制部 2 控制输入部 1、文本缓冲器 4、显示部 3、辞典选择、顺序指定部 8、假名汉字变换控制部 5 等。

以下对上述构成的具有专业辞典转换功能的汉字变换装置,说

明其工作。使用者根据编制文件内容,用辞典选择、顺序指定部 8 从基本辞典 70 及多个专业辞典 70-7n 中选择将成为检索对象的、汉字与单词登记可能性高的多个辞典,并再指定其检索顺序,将其结果登记于辞典群保持部 91、92。例如,编制有关信息处理文件时,把通常变换时想检索的辞典(基本辞典及例如对应文件内容的信息处理用语辞典)、欲变换为固有名词等时把想检索的辞典(例如,姓名辞典、地名辞典、基本辞典)分别以其优先顺序预先登记在辞典群保持部 91、辞典群保持部 92 中。由输入部 1 输入的假名字符串保持在文本缓冲器 4 内并由显示部 3 显示。假名汉字处理中,假名汉字变换控制部 5 一边用辞典检索部 6 检索登记于辞典群保持部 91、92 中的辞典,一边将假名字符串变换成假名汉字混和的字符串。进而用变换后的字符串置换文本缓冲器 4 内的假名字符串,显示在显示部 3 上。此时,在预先将上述辞典登记在辞典群保持部 91、92 中的情况下,通常的假名汉字变换时,因按登记顺序检索登记在辞典群保持部 91 中的辞典,信息处理用语比一般用语先输出。若文件中出现人名等固有名词,则操作输入部 1 的辞典群转换键,使被检索的辞典群转换至登记在辞典群保持部 92 中的辞典。若再次操作辞典的转换键,则回到辞典群保持部 91。就这样通过操作辞典群转换键来进行辞典群的转换。

另外,因有些内容与本发明宗旨无直接关系,省去有关图示与说明。例如同一发音下有多个汉字串、词语的情况下用于将它们按优先顺序或一次全显示于屏幕后让使用者特别指定想要的字语的程序及其所必要的结构;用于将使用者特别指定的汉字串显示于眼下正写作中的文章的所定位置的程序及结构;用于将由许多输入标音字符组成的标音字符串划分成作为整体的汉字变换对象的必要程序,例如最长一致法及实现其的结构,与通常的中文文字处理器、日文文字处理器相同或基本相同。

然而,上述以往技术装置存在以下问题。

第一个问题:由工作人员按输入文件内容选择辞典,且必需转换

专业辞典。而正确转换专业辞典,对于使用者来说成为大负担。还有,登记在辞典群保持部内的辞典因专业领域不正确时,由最长一致法等截出的作为汉字变换对象的词态素读音码(标音字符串)以及文字变换就不正确,汉字变换效率变差。

第二问题:若输入文件中包括多个专业的单词,为此而必需对专业选定作繁琐变更时,对使用者来说,转换专业辞典很费时间。

本发明的目的在于解决上述问题,提供一种能自动又确切选定专业辞典的汉字变换装置。

为实现上述目的,本发明第一形态的汉字变换装置,是具有将与输入标音字符串对应的词态素读音码截出后送至辞典检索部的输入部、由基本辞典与多本专业辞典构成的辞典部和显示部,其特征是还具有:①对各专业辞典记录其使用次数的辞典记录部;②让上述辞典使用记录部所记录的使用次数多的优先来决定各辞典检索顺序的辞典选择部;③以上述输入部送来的读音记号为检索关键码,按上述辞典选择部决定的检索顺序进行各辞典的检索,取出与读音记号对应的所有同音字、候选词以及登记有它们的专业辞典的编码,再送至字符变换部的辞典检索部;④使上述辞典检索部检出的同音字、候选词显示于显示部,让使用者确认正确的变换,同时将登记有被认为正确的同音字、词的辞典的编码送至辞典管理部的字符变换部;⑤根据上述字符变换部送来的辞典编码,按所定程序更新上述辞典使用记录部记录的各专业辞典的使用次数的辞典管理部。

本发明第二形态的汉字变换装置,是在第一形态基础上再加上这样的特征,即上述辞典管理部更新上述辞典使用记录部的记录时,在记录有使用次数的多本专业辞典中,规定数量专业辞典的使用次数达到预定阈值情况下,对迄今的或者汇总汉字变换结束前的使用次数未达到其阈值的专业辞典、将其使用次数减去规定值而为0,或者迄今使用次数为0而经上述减法结果为负的专业辞典,设定为不使用。

本发明第三形态的汉字变换装置,是在第一或二形态基础上再加上这样的特征,即上述词典检索部在规定数量词典中检索到与读音记号对应的同音字、候选词时,上述显示部只将该阶段中检索到的同音字、候选词显示于显示部,以便使用者确认正确的变换。

图1是本发明汉字变换装置一实施例的结构图;

图2是图1实施例的处理流程图;

图3是上述实施例的词典管理部的工作流程图;

图4从概念上表示图1实施例的词典使用记录部的内容;

图5是以往汉字变换装置的结构图。

图中有关标号含义为以下。1:输入部、2:控制部、3:显示部、4:文本缓冲部、5:字符变换部、6:词典检索部、70:基本词典、71……7n:专业词典、81:词典管理部、82:词典使用记录部、90:词典选择部。

以下举实施列说明本发明。

(第1实施例)

本实施例是关于本发明第一形态和第二形态的发明。

图1是本发明汉字变换装置一实施例的结构图。图中,1是输入部、2是控制部、3是显示部、4是文本缓冲部、5是文字变换部、6是词典检索部、70是基本词典、71-7n是专业词典、81是词典管理部、82是词典使用记录部、90是词典选择部。控制部2、显示部3、文本缓冲部4、基本词典70及专业词典71-7n的结构、原理与以往技术相同或没有大的区别,这方面的说明只要与本发明无关,就省略或是简单说明。基本词典70收存了以读音记号为检索关键码的对应的基本单词。词典使用记录部82以词典编码为检索关键码,记录使用次数值或不使用记号“X”的值,而且按其使用次数值大的比小的排前的顺序进行记录。专业词典71-7n,对多个专业领域收集各专业固有的单词,并以读音记号为检索关键码收存该专业中对应的单词。词典选择部90参考词典使用记录部82所记录的各专业词典的使用次数及排列顺序,决定使用的专业词典及

其检索顺序。还根据辞典管理部 81 的指示, 决定作为检索对象的专业辞典和检索顺序。辞典检索部 6, 以读音记号为检索关键码, 根据辞典选择部 90 指定的专业辞典及其检索顺序, 取出对应的所有同音字、候选词。辞典管理部 81, 对每个由使用者确认为正确变换的字符串中的汉字、单词, 根据它们的辞典编码更新处理辞典使用记录部 82 所记录各专业辞典的使用次数记录及排列顺序。字符变换部 5, 通过辞典检索部 6 将输入标音字符串变换成候选变换的汉字串、单词后, 使之显示于显示部 3 上, 然后, 让使用者选择想要的汉字串、单词, 同时将登记有确认为正确的汉字串、单词的辞典的编码送至辞典管理部 81。显示部 3 依次显示辞典检索部 61 检索过的同音字、候选词, 进而将使用者认为正确的汉字串、单词显示于写作中的文章的所定位置。

图 2 和图 3 是本实施例的工作流程图。以下根据图 2 说明其基本的工作流程。

(S21): 使用者通过输入部将作为汇集一起的汉字变换对象的标音字符串进行输入。并在后文将举的具体例中, 输入构成 7 个音节的标音字符串。

(S22): 输入部从输入标音字符串中截出与成为一体变换对象的单词对应的词态素读音码。并在后文将举的具体例中, 2 个或者 3 个音节构成单词。

(S23): 辞典检索部参照辞典使用记录部记录的使用次数, 检索专业辞典或基本辞典的内容, 显示其对应的字、候选词。此时, 若是检索被认为是未变换或错变换的词态素的全部读音码, 则对该词态素进行字符变换。图 4 从概念上表示该辞典使用记录部的使用次数的记录内容。

(S24): 使用者参见显示部, 判断被变换的汉字是否全都正确。若有不正确的汉字, 则应将它指出, 并参照该词态素或者是包括其在内的成为一体变换对象的词态素的读音码再次进行检索, 回到 (S23)。若没有, 则进入 (S25)。

(S25): 参照包含被认为是正确对应词态素读音码的汉字、单词的辞典的辞典编码, 更新辞典使用记录部的内容。后文将详细说明其更新时的详细工作过程。

(S26): 将正确变换的汉字串显示于写作中的文章的所定位置上。

下面, 根据图 3 说明辞典管理部的工作过程。

(S301): 设定阈值 TH 以及作为汇总的汉字变换对象而输入的标音字符串中所含的成一体性汉字变换的单词的总数 h。

(S302): 将与确定为正确变换的单词对应的汉字、单词的辞典编码依次存入暂存装置。为此, 设定变量 A[J]。其中, $J=1, 2, \dots, h$, 各 A[J] 的内容为辞典编码。

(S303): 设变量 J 的起始值为 1。

(S304): 判断是否 $J > h$ 。若为否定, 则进入 (S305)。

(S305): 以 A[J] 之值为检索关键码, 参见辞典使用记录部, 取出该专业辞典的使用次数, 作为变量 CNT 进行存贮。

又因 $CNT=X$ 情况下, 该辞典为不使用, 所以不会化费太多的检索时间。

(S306): CNT 加 1 后, 就进入上述 (S307)。

(S307): 判断是否 $CNT > TH$ 。若为肯定, 则为了转成更新处理, 进入 (S308), 若为否定, 则进入 (S310)。

(S308): 让当前辞典使用记录部所记录的专业辞典的使用次数减去规定值 D。然后, 进入 (S309)。

(S309): 将此结果设定为新使用次数。使用次数为 0 或者为负数的专业辞典, 将其使用次数设定为 X。进入 (S310)。

(S310): 以 A[J] 为检索关键码求该专业辞典, 并将该新的 CNT 之值记录于辞典使用记录部。

(S311): J 加上 1 后, 返回至应以下一个汉字、单词为对象而进行相同处理的 (S304)。

(S312): 根据使用次数, 结束辞典使用记录部的使用次数的记录

和排列顺序的更新。

图 4 从概念上表示辞典使用存贮部的存贮内容。本图中, 第 1 列表示专业辞典编码, 第 2 列表示使用次数, 将它们按使用次数大的顺序从上至下进行排列。图中, 使用次数为 X 的表示该专业辞典定为不使用。

本实施例中, 基本辞典除作为本发明的作用对象外, 还被列为第一优先检索, 将辞典管理部的阈值和比较专业辞典使用次数的规定数设定为 1。

实际上, 因基本辞典的使用频度比任何一个专业辞典高, 这与将基本辞典也作为本发明的作用对象后设置阈值与比较使用次数的辞典的规定值为 2 实质上是相同的。

但是, 由于此情况下只是基本辞典的使用频度变高, 往往总是其使用次数的记录数变高, 因此有必要采取设置其上限等措施。

接着, 举图 4 所示辞典使用存贮部为例, 按图 3 所示辞典管理部的工作过程, 具体说明其如何变化。

现若输入“lian2jyuen1zuen1shou3guo2ji4fa3”这样的标音字符串, 则根据使构成音节数第 1 优先、使输入顺序第 2 优先而截出作为汉字变换对象的音节的最长一致法截出词态素的读音码, 结果为“lian2jyuen1 zuen1shou3 guo2ji4fa3”。然后, 参见专业辞典和基本辞典, 将上述标音字符串变换成“联军遵守国际法”这样的字符串。其关连的辞典编码分别是“联军”为 5(军事)、“遵守”为 0(基本辞典)、“国际法”为 7(法律)。又因为在前面图 3 所示工作过程说明中, 各 A[J] 的内容只是“专业辞典的编码”, 专业辞典的种类全部是 9 个, 因此(S302)中的 A[J] 之值为 1~9。

接着, 说明图 3 的辞典管理部的处理过程。

(S301): 由于上述一连串标音字符串的变迁中, 成为一体性汉字变换对象的汉字串、单词的个数为 3, 故确定 $h=3$ 。为便于说明, 在此将 TH 的值设定为 137。

(S302): $A[1]=5$ 、 $A[3]=7$

(S303): 设变量 J 为 1。

(S304): 因 $J < 3 = h$, 进入 (S305)。

(S305): 以 A [1] 之值 5 为检索关键码, 从辞典使用记录部取出其使用次数 137 后存入 CNT 中。

(S306): CNT 加上 1, 成为 138。

(S307): 因 $CNT > TH$, 进入 (S309)。

(S308): 辞典使用记录部由于减去规定值 10, 更新成以下那样。(但只表示图 4 中的更新部分。而且, 表示以后更新内容时也相同。)

5(军事)138

3(经济)59

7(法律)1

8(会计)X

在以上条件下, 进入 (S310)。

(S310): 以 A [1] 之值 5 为检索关键码求出专业辞典, 并将当前的 CNT 之值作为使用次数, 记录于辞典使用记录部。

(S311): J 加上 1 为 2 后, 返回上述 (S304)。

J=2、J=3 中也进行相同处理工作。于是, 此状态下的辞典使用记录部的内容依次表示为如下。

J=2:

5(军事)139

3(经济)59

7(法律)1

8(会计)X

J=3:

5(军事)139

3(经济)59

7(法律)2

8(会计)X

此时, 因(304)中 $J=4 > "3"$, 进入(S312)。

(S312): 结束使用次数的记录及与其对应的排列顺序的变更处理。此时, 最后使用的专业辞典与其排列顺序相同。

(第 2 实施例)

本实施例, 除了在前述第 1 实施例中, 当某专业辞典的使用次数达到预定的阈值时, 辞典管理部将达到此情况下的使用次数改设为 0 以外, 都与第 1 实施例相同。

因此, 省去的各部结构图的图示以及以此为基础的工作作用说明等, 并省去处理过程图中相同的工作说明, 只说明本实施例固有的工作, 处理过程等。

在前述的(S307)中, 辞典管理部将有的专业辞典之值设为 0。于是, 辞典使用记录部的内容为以下那样。

5(军事)0

3(经济)0

7(法律)0

8(会计)1

同样, 在(S311)中, 于 $J=2$ 的工作结束阶段, 辞典使用记录部的内容为以下那样。

5(军事)1

3(经济)0

7(法律)0

8(会计)X

在 $J=3$ 的工作结束阶段, 同样为以下那样。

5(军事)1

3(经济)0

7(法律)1

8(会计)X

在上述使用次数下, 辞典选择部变更各专业辞典的排列顺序。因此, 最后为下述那样。

5(军事)1

7(经济)1

3(法律)0

8(会计)X

与前面第 1 实施例相比,因本实施例将所有的专业辞典的使用次数改设为 0,这对作为起初使用的专业辞典设定数(种类)为多的情况变得有利。

(第 3 实施例)

本实施例是关于本发明的第三形态。

本实施例,因基本结构与前述第 1 实施例相同,省去结构图等。

不同之处只是辞典选择部内装计数部,该计数部每逢一检索关键词即读音记号就计算作为检索对象的辞典的个数,若该计算值在预定的规定值内并检索成功,则将该情况通知字符变换部。而且,接到该通知的字符变换部使该阶段检索到的同音字、候选词显示于显示部的所定位置,以便使用者确认正确的选择。并且,具有将通过检索其后优先序位低的辞典检出的同音词、候选词不显示于显示部而进行保存的暂存部。

至于计数部、根据来自计数部的通知使字符变换部在中途阶段发挥作用的扳机信号(触发信号)发生部、暂存部的结构都因是众所周知的技术,省去其内容说明。

与前面实施例相比,本实施例在检索中途阶段单把正确率高的同音字、候选词作为使用者想要的汉字、单词的确认对象而进行显示、选择,从而提高了汉字变换速度。

以上根据实施例说明了本实施例,不用说,本发明不只限于上述实施例。也就是说,例如以下情况也包含于本发明内。

(1)第 1、第 2 实施例中,是加 X 表示不使用的专业辞典,但也可以采用其它任意标记、措施。同样,各专业辞典的使用次数减去规定值,不是逐个单词(具体例中的 A[J],对 J=1、2、3 逐个地)进行,而是对作为汇总的汉字变换对象的全部单词的各专业辞典添

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☒ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.